VOLUME 29 · NUMBER 1 · JANUARY 1 2011

JOURNAL OF CLINICAL ONCOLOGY

CORRESPONDENCE

Breast Carcinomas Arising at a Young Age: Unique Biology or a Surrogate for Aggressive Intrinsic Subtypes?

TO THE EDITOR: In the July 10, 2008 issue of *Journal of Clinical Oncology*,[1] we reported a large-scale genomic analysis illustrating that mRNA expression levels of key breast cancer-associated genes, *ER-α*, *ER-β*, epithelial growth factor receptor (*EGFR*), and human epidermal growth factor receptor 2 (*HER2*) occurred in an age-related manner. Moreover, when stratified by age, breast tumors arising in younger women (≤ 45 years) were enriched with > 350 biologically relevant gene sets compared with those arising in older women (≥ 65 years).[2] Breast cancer is no longer viewed as a single disease, but rather a compilation of several distinct subtypes defined via gene expression analysis.[2] Microarray techniques have divided breast cancer into intrinsic subtypes: luminal A, luminal B, HER2-enriched, and basal-like, each with unique prognostic and therapeutic implications.[3,4] On the basis of findings from Carey et al,[5] we hypothesized that (1) breast tumors arising in younger women may be more enriched for aggressive subtypes and (2) age-specific biologic differences observed in breast carcinomas may be highly subtype dependent. To evaluate the relationship between age and breast cancer subtype, and to account for potential confounding variables not previously included, we have performed new analyses on data from Anders et al[1] and include a similar analysis performed on a second independent microarray-based breast tumor data set.

To explore our hypotheses, we chose to reanalyze our previous data set; however, we limited our current analyses to a combination of two of the four large data sets used previously, genomic spatial event (GSE) 4922[6] and GSE7849,[7] termed data set A. Please note that two of the four previous data sets were excluded (GSE2034[8] and GSE3143[9]) because they lacked complete clinical data (ie, histologic grade). Our first goal was to define the distribution of intrinsic subtypes assigned by the PAM50[10] to determine whether subtype correlated with age distinction. We hypothesized that more aggressive subtypes (ie, basal-like) would be over-represented among breast carcinomas arising in younger women (≤ 45 years), whereas older women (≥ 65 years) would more commonly be diagnosed with luminal tumors.[5] As expected, there was a significant association between subtype and age (P = 3.8e-06; Table 1). Specifically, a higher proportion of younger women were diagnosed with basal-like (odds ratio [OR], 12.27; 95% CI, 3.96 to 45.0) and HER2-enriched (OR, 4.63; 95% CI, 1.50 to 16.48) breast tumors.

Recognizing that age-specific differences in subtype distribution were present, we next examined other potential confounding variables and noted that grade and sample source were associated with age (Table 1). We hypothesized that accounting for all significant

Table 1. Age-Specific Clinical Characteristics of Data Set A (GSE4922 and GSE7849; age ≤ 45 and ≥ 65 years)

	All (N = 192)		Younger (≤ 45; n = 48)		Older (≥ 65; n = 144)				
Characteristic	No.	%	No.	%	No.	%	Odds Ratio	*P*	95% CI
Estrogen receptor									
Positive	156	83	38	81	118	84	1.28	.58	0.51 to 2.96
Negative	31	17	9	19	22	16	1.00	NA	NA
Nodal status									
Positive	57	31	17	35	40	30	0.77	.46	0.38 to 1.57
Negative	126	69	31	65	95	70	1.00	NA	NA
Tumor size									
≤ 2 cm	109	57	23	48	86	60	1.00	NA	NA
> 2 cm	83	43	25	52	58	40	0.62	.16	0.32 to 1.20
Histologic grade									
1	41	22	6	13	35	25	0.31	.015	0.10 to 0.80
2	82	44	16	36	66	47	0.43	.025	0.20 to 0.90
3	63	34	23	51	40	28	1.00	NA	NA
Subtype (PAM50)									
Lum A	51	27	5	13	46	38	1.00	NA	NA
Lum B	46	24	5	13	41	34	1.12	.87	0.28 to 4.45
HER2	35	18	12	31	23	19	4.63	7.0e-03	1.50 to 16.48
Basal	29	15	17	44	12	10	12.27	5.2e-06	3.96 to 45.0
Normal	31	16	9	19	22	15	3.65	.033	1.11 to 13.50
Source									
GSE4922	156	81	33	69	123	85	1.00	NA	NA
GSE7849	36	19	15	31	21	15	0.38	.015	0.17 to 0.83

Abbreviations: NA, not available; PAM, prediction analysis of microarray; Lum, luminal; HER2, human epidermal growth factor receptor 2; GSE, genomic spatial event.

Table 2. Age-Defined Gene Expression Differences (q < 0.05) Between Breast Carcinomas Arising in Younger (age ≤ 45 years) Versus Older (age ≥ 65 years) Patients Are Minimal After Correction for Significant Clinicopathologic Features

	Number of Genes Differentially Expressed	
Data Set	Uncorrected	Corrected
A*	693	0
B†	2,154	1

*Data set A corrected for subtype, grade, and data set source.
†Data set B corrected for subtype, estrogen receptor status, nodal status, and grade.

clinicopathologic features, namely intrinsic subtype and grade, could have an effect on the number of age-associated genes, thus we used a statistical model that can account for confounding variables. To test our hypothesis, we built a linear regression model for each gene's expression value using age alone or in combination with significant clinical variables, (ols function in R package Design). Before analysis, the $\log_2$ intensities of the gene expression from data set A, Affymetrix one-channel data, were row (gene) median centered, and column (sample) standardized. The two data sets (GSE4922[6] and GSE7849[7]) were combined, using distance weighted discrimination to detect and remove batch bias.[11] A linear model of gene expression was defined by age (unadjusted model). A second linear model contained additive terms for age, grade, subtype, and sample source (adjusted model). Higher order interactions were not considered. We then transformed the *P* values for the age term to q values (the false discovery rate at which the gene is significant) using the method of Benjamini and Hotchterg.[12] A false discovery rate of 5% was used to identify significant genes. Within data set A, the unadjusted model of breast tumor gene expression differences by age alone (≤ 45 *v* ≥ 65 years) yielded 693 genes differentially expressed (q < 0.05; Table 2). Correction for the significant clinicopathologic features (grade, subtype, sample source; Table 1) with the adjusted model yielded zero gene differences (q < 0.05) between breast tumors of previously defined age groups. Recognizing that gene differences diminished to zero, we did not believe gene set enrichment analysis as previously reported would have added further to this analysis.

As is standard for the field, we elected to evaluate our findings in a second independent data set. To conduct our secondary analysis, we pooled 344 clinically annotated breast tumors assayed on the full genome Agilent microarrays from four publications[10,13-15] and 12 new arrays (GSE20624, obtained with institutional review board approval), by selecting the subset of samples from these publications that had the same complete clinical data as data set A; this was termed data

Table 3. Age-Specific Clinical Characteristics of Data Set B (age ≤ 45 and ≥ 65 years)

	All (N = 200)		Younger (≤ 45; n = 92)		Older (≥ 65; n = 108)				
Characteristic	No.	%	No.	%	No.	%	Odds Ratio	*P*	95% CI
Estrogen receptor									
Positive	111	57	38	42	73	69	0.33	1.9e-4	0.18 to 0.60
Negative	85	43	52	58	33	31	1.00	NA	NA
Nodal Status									
Negative	83	43	29	32	54	51	1.00	NA	NA
Positive	112	57	61	68	51	49	2.21	7.2e-3	1.24 to 4.02
Tumor size									
≤ 2 cm	43	22	18	20	25	24	1.00	NA	NA
> 2 cm	152	78	71	80	81	76	1.21	.58	0.61 to 2.44
Grade									
1	12	6	6	7	6	6	0.87	.82	0.25 to 2.99
2	55	29	13	15	42	40	0.27	1.8e-4	0.13 to 0.54
3	125	65	67	78	58	55	1.00	NA	NA
Race									
African American	77	39	41	45	36	33	1.00	NA	NA
White	123	62	51	55	72	67	0.62	.11	0.35 to 1.11
Subtype (PAM50)									
Lum A	66	33	19	22	47	48	1.00	NA	NA
Lum B	34	17	13	15	21	21	1.53	.35	0.63 to 3.69
HER2	23	12	13	15	10	10	3.16	.022	1.18 to 8.73
Basal	62	31	42	48	20	20	5.10	1.1e-5	2.43 to 11.11
Normal	15	8	5	5	10	9	1.25	.72	0.34 to 4.10
Source									
Thomas Jefferson University	15	8	9	10	6	6	1.00	NA	NA
University of Chicago	28	14	13	14	15	14	1.70	.42	0.47 to 6.49
UNC-CH	71	36	39	42	32	30	1.22	.74	0.39 to 4.07
University of Utah	17	9	6	7	11	10	2.64	.19	0.63 to 12.10
Washington University	69	35	25	27	44	41	2.59	.10	0.82 to 8.74

Abbreviations: NA, not available; PAM, prediction analysis of microarray; Lum, luminal; HER2, human epidermal growth factor receptor 2; UNC-CH, University of North Carolina at Chapel Hill.

set B. Similar to data set A, there was an association between age (≤ 45 $v \geq 65$ years) and intrinsic subtype in data set b ($P = 1.6e-4$, Table 3). Younger women were more commonly diagnosed with basal-like breast tumors (OR, 5.1; 95% CI, 2.43 to 11.11) and HER2-enriched breast tumors (OR, 3.16; 95% CI, 1.18 to 8.73). We used the same statistical approach described above to evaluate age-specific gene expression differences. Specific to data set B, gene expression data is Agilent two-channel data. The $\log_2$(R/G) of the gene expression was LOWESS normalized, row (gene) median centered, and column (sample) standardized. Comparison of breast tumor gene expression differences by age alone (≤ 45 $v \geq 65$) yielded 2,154 genes differentially expressed between age-defined classes ($q < 0.05$; Table 2). Correction for additive effects of significant clinicopathologic features (including intrinsic subtype, estrogen receptor status, grade, and nodal status, Table 3) yielded only one gene difference between breast tumors of age-defined groups, (*SLC25A20*, Solute carrier family 25). An identical analysis evaluating age-specific gene expression differences by age less than 45 versus ≥ 45 years yielded identical findings; within data set B, 778 genes differentially expressed by age group diminished to zero gene differences when correcting for subtype and other significant clinicopathologic features (ER status and histologic grade) that differed between age-defined groups.

Results of this analysis continue to refine our understanding of the biology of breast cancer arising in younger women and support our hypothesis that younger women's breast tumors are enriched for more aggressive intrinsic subtypes, namely, basal-like. This finding is complementary to our prior report illustrating breast tumors arising in younger women are characterized by lower mRNA expression of *ER-α, ER-β,* and *PR,* but higher expression of *EGFR*[1], a known marker of the basal-like subtype.[16] Although we recognize that our current analysis is not population-based, our results are entirely consistent with those of the population-based Carolina Breast Cancer Study, which reported that basal-like breast tumors occurred at a higher prevalence among premenopausal African American patients.[5] Most important, our current analysis strongly suggests that biologic differences present between breast carcinomas arising at the extremes of age are strongly influenced by genes associated with intrinsic breast cancer subtype and grade, both of which are highly correlated with age. We recognize that our analysis is not designed to address age-related differences in tumor-host interfaces (which most certainly vary by age) and may not be entirely reflective of tumors arising in very young women (< 35 years), both areas deserving of future research in (ideally) prospectively collected, clinically annotated data sets. Our results, however, suggest there are few age-specific differences in breast tumor biology. Age alone does not appear to provide an additional layer of biologic complexity above that of breast cancer subtype and grade; therefore, when considering treatment programs, decisions should be driven by subtype biology and performance status, and much less influenced by age.

Carey K. Anders, Cheng Fan, Joel S. Parker, and Lisa A. Carey
Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC

Kimberly L. Blackwell
Duke University Medical Center, Durham, NC

Nancy Klauber-DeMore and Charles M. Perou
Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC

ACKNOWLEDGEMENT

Supported by National Institutes of Health Grant No. 5K12CA120780-03 (C.A.) and National Cancer Institute Breast SPORE program No. P50-CA58223-09A1 (University of North Carolina at Chapel Hill).

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Although all authors completed the disclosure declaration, the following author(s) indicated a financial or other interest that is relevant to the subject matter under consideration in this article. Certain relationships marked with a "U" are those for which no compensation was received; those relationships marked with a "C" were compensated. For a detailed description of the disclosure categories, or for more information about ASCO's conflict of interest policy, please refer to the Author Disclosure Declaration and the Disclosures of Potential Conflicts of Interest section in Information for Contributors.

Employment or Leadership Position: Charles M. Perou, University Genomics (U) **Consultant or Advisory Role:** None **Stock Ownership:** Charles M. Perou, Bioclassifier **Honoraria:** None **Research Funding:** None **Expert Testimony:** None **Other Remuneration:** None

REFERENCES

1. Anders CK, Hsu DS, Broadwater G, et al: Young age at diagnosis correlates with worse prognosis and defines a subset of breast cancers with shared patterns of gene expression. J Clin Oncol 26:3324-3330, 2008

2. Perou CM, Sørlie R, Eisen MB, et al: Molecular portraits of human breast tumors. Nature 406:747-752, 2000

3. Sørlie R, Perou CM, Tibshirani R, et al: Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implication. Proc Natl Acad Sci U S A 98:10869-10874, 2001

4. Sørlie R, Tibshirani R, Parker J, et al: Repeated observation of breast tumor subtypes in independent gene expression datasets. Proc Natl Acad Sci U S A 100:8418-8423, 2003

5. Carey LA, Perou CM, Livasy CA, et al: Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. JAMA 295:2492-2502, 2006

6. Ivshina AV, George J, Senko O, et al: Genetic reclassification of histologic grade delineates new clinical subtypes of breast cancer. Cancer Res 66:10292-10301, 2006

7. Anders C, Acharya CR, Hsu DS, et al: Age-specific differences in oncogenic pathway deregulation seen in human breast tumors. PLoS One 3:e1373, 2008

8. Wang Y, Klijn JG, Zhang Y, et al: Gene expression profiles to predict distant metastasis in lymph-node negative primary breast cancer. Lancet 365:671-679, 2005

9. Bild AH, Yao G, Chang JT, et al: Oncogenic pathway signatures in human cancers as a guide to targeted therapies. Nature 439:353-357, 2006

10. Parker JS, Mullins M, Cheang MC, et al: Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol 27:1160-1167, 2009

11. Benito M, Parker J, Du Q, et al: Adjustment of systematic microarray data biases. Bioinformatics 20:105-114, 2004

12. Yoav B and Yosef H: Controlling the false discovery rate: A practical and powerful approach to multiple testing. J Royal Stat Soc, Series B (Methodological) 57:289-300, 1995. http://www.istor.org/stable/2346101

13. Hoadley KA, Weigman VJ, Fan C, et al: EGFR associated expression profiles vary with breast tumor subtype. BMC Genomics 8:258, 2007

14. Hu Z, Fan C, Oh DS, et al: The molecular portraits of breast tumors are conserved across microarray platforms. BMC Genomics 7:96, 2006

15. Hu Z, Fan C, Livasy C, et al: A compact VEGF signature associated with distant metastases and poor outcomes. BMC Med 7:9, 2009

16. Livasy CA, Karaca G, Nanda R, et al: Phenotypic evaluation of the basal-like subtype of invasive breast carcinoma. Mod Pathol 19 264-71, 2006

DOI: 10.1200/JCO.2010.28.9199; published online ahead of print at www.jco.org on November 29, 2010

